

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-172697

(P2000-172697A)

(43) 公開日 平成12年6月23日(2000.6.23)

(51) Int.Cl.⁷

G 0 6 F 17/30

識別記号

F I

C 0 6 F 15/403

テーマコード(参考)

3 4 0 A 5 B 0 7 6
3 6 0 C

審査請求 未請求 請求項の数25 O L (全 14 頁)

(21) 出願番号

特願平10-344357

(22) 出願日

平成10年12月3日(1998.12.3)

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 藤原 義仁

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

(72) 発明者 安藤 尚隆

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

(74) 代理人 10006/736

弁理士 小池 晃 (外2名)

Fターム(参考) 5B075 ND03 NK46 PQ02 PQ46 PR06
QM08

(54) 【発明の名称】 顧客情報検索方法及び顧客情報検索装置、データ作成方法、並びにデータベース

(57) 【要約】

【課題】 顧客の規則抽出の精度を向上させることができる。

【解決手段】 顧客データベースDB₁は、顧客の価値観等を含む心理側面に関する顧客内的情報と、少なくとも顧客の性別及び年齢等からなる基本属性を含む顧客外的情報とが蓄積されて形成されている。

プロファイル 項目	デモグラフィック情報				電話/サービス 履歴・利用履歴				人物データ			
	氏名	性別	年齢	住所コード	電話番号	年次コード	高年次コード	年齢コード	年次コード	高年次コード	年齢コード	年次コード
顧客ID	101.000.001	山田	25	44	27	1	35	2	3	3	1999	4.01
	100.000.002	田中	28	78	32	1	24	4	4	7	1999	4.03
	100.000.003	山下	11	12	14	2	47	1	6	7	1999	4.15
	100.000.004	鈴木	78	47	41	1	18	3	0	2	1999	5.17
	100.000.005	西村	62	45	22	2	33	3	2	1	1999	5.08
	109.999.999	山田	13	83	27	2	42	4	3	7	2000	1.17
	109.999.999	山田	29	98	13	1	51	2	8	7	2001	8.27
	110.000.000	川崎	65	55	12	1	21	1	2	6	2001	9.12

DB₁ 顧客データベース

(2) 000-172697 (P2000-172697A)

【特許請求の範囲】

【請求項1】 顧客に関して蓄積された情報に基づいて、所望の顧客情報を検索する顧客情報検索方法であって、顧客の心理側面に関する顧客内的情報と、少なくとも顧客の基本属性を含む顧客外的情報とを記憶手段に蓄積し、

上記記憶手段に蓄積した顧客内的情報及び顧客外的情報に基づいて、顧客の特定の商品又はサービスに対する有望度を算出することを特徴とする顧客情報検索方法。

【請求項2】 上記有望度に基づいて顧客を順位付けすることを特徴とする請求項1記載の顧客情報検索方法。

【請求項3】 上記顧客内的情報は、顧客の価値観を少なくとも含む心的側面に関する情報であることを特徴とする請求項1記載の顧客情報検索方法。

【請求項4】 上記顧客外的情報には、顧客の購買履歴を含む行動履歴に関する情報も含まれており、上記基本属性には、少なくとも顧客の性別及び年齢が含まれていることを特徴とする請求項1記載の顧客情報検索方法。

【請求項5】 上記顧客内的情報及び上記顧客外的情報に基づいて作成した計算式により上記有望度を算出することを特徴とする請求項1記載の顧客情報検索方法。

【請求項6】 顧客に関して蓄積された情報に基づいて、所望の顧客情報を検索する顧客情報検索装置であって、顧客の心理側面に関する顧客内的情報と、少なくとも顧客の基本属性を含む顧客外的情報とが蓄積される記憶手段と、

上記記憶手段に蓄積した顧客内的情報及び顧客外的情報に基づいて、顧客の特定の商品又はサービスに対する有望度を算出する有望度算出手段とを備えることを特徴とする顧客情報検索装置。

【請求項7】 上記有望度に基づいて顧客を順位付けすることを特徴とする請求項6記載の顧客情報検索装置。

【請求項8】 上記顧客内的情報は、顧客の価値観を少なくとも含む心的側面に関する情報であることを特徴とする請求項6記載の顧客情報検索装置。

【請求項9】 上記顧客外的情報には、顧客の購買履歴を含む行動履歴に関する情報も含まれており、上記基本属性には、少なくとも顧客の性別及び年齢が含まれていることを特徴とする請求項6記載の顧客情報検索装置。

【請求項10】 上記顧客内的情報及び上記顧客外的情報に基づいて作成した計算式により上記有望度を算出することを特徴とする請求項6記載の顧客情報検索装置。

【請求項11】 顧客の心理側面に関する顧客内的情報と、少なくとも顧客の基本属性を含む顧客外的情報とを記憶手段に蓄積することを特徴とするデータ作成方法。

【請求項12】 上記顧客内的情報は、顧客の価値観を

少なくとも含む心的側面に関する情報であることを特徴とする請求項11記載のデータ作成方法。

【請求項13】 上記顧客外的情報には、顧客の購買履歴を含む行動履歴に関する情報も含まれており、上記基本属性には、少なくとも顧客の性別及び年齢が含まれていることを特徴とする請求項11記載のデータ作成方法。

【請求項14】 顧客に関する情報から構成される既存の顧客データベースから商品又はサービスの有望度を得るために取り出した第1の顧客情報と、商品又はサービスの有望度を得るために新たに取得した顧客に関する情報を示す第2の顧客情報とを有し、上記第1の顧客情報と上記第2の顧客情報の間に共通して存在する規則が抽出されることを特徴とするデータベース。

【請求項15】 上記規則に基づいて、潜在顧客の特定の商品又はサービスに対する有望度を得ることを特徴とする請求項14記載のデータベース。

【請求項16】 上記有望度に基づいて顧客を順位付けすることを特徴とする請求項15記載のデータベース。

【請求項17】 上記第1の顧客情報及び／又は上記第2の顧客情報は、顧客の心理側面に関する顧客内的情報及び／又は少なくとも顧客の基本属性を含む顧客外的情報であることを特徴とする請求項14記載のデータベース。

【請求項18】 上記顧客内的情報は、顧客の価値観を少なくとも含む心的側面に関する情報であることを特徴とする請求項17記載のデータベース。

【請求項19】 上記基本属性には、少なくとも顧客の性別及び年齢が含まれていることを特徴とする請求項17記載のデータベース。

【請求項20】 顧客に関する情報から構成される既存の顧客データベースから商品又はサービスの有望度を得るために取り出した第1の顧客情報と商品又はサービスの有望度を得るために新たに取得した顧客に関する情報を示す第2の顧客情報の間に共通して存在する規則を抽出することを特徴とする顧客情報検索方法。

【請求項21】 上記規則に基づいて、潜在顧客の特定の商品又はサービスに対する有望度を得ることを特徴とする請求項20記載の顧客情報検索方法。

【請求項22】 上記有望度に基づいて顧客を順位付けすることを特徴とする請求項21記載の顧客情報検索方法。

【請求項23】 上記第1の顧客情報及び／又は上記第2の顧客情報は、顧客の心理側面に関する顧客内的情報及び／又は少なくとも顧客の基本属性を含む顧客外的情報を取得することを特徴とする請求項20記載の顧客情報検索方法。

【請求項24】 上記顧客内的情報は、顧客の価値観を少なくとも含む心的側面に関する情報であることを特徴

(3) 000-172697 (P2000-172697A)

とする請求項23記載の顧客情報検索方法。

【請求項25】 上記基本属性には、少なくとも顧客の性別及び年齢が含まれていることを特徴とする請求項23記載の顧客情報検索方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、顧客に関する情報に基づいて、所望の顧客を検索する顧客情報検索方法及び顧客情報検索装置、データを作成するデータ作成方法、並びに各種情報からなるデータベースに関し、詳しくは、顧客の性別及び年齢等のプロフィールに基づいて所望の顧客情報を検索するのに好適な顧客情報検索方法及び顧客情報検索装置、データ作成方法、並びにデータベースに関する。

【0002】

【従来の技術】消費者の将来における消費動向を探る一手段として、顧客データベース化がある。従来の顧客データベースに保存されている項目は、性別、年齢、職業といった顧客の基本属性や商品の購買行動の時系列記録といった行動履歴からなっていた。

【0003】このような顧客データベースにより、データを統計的に処理することで顧客の行動パターンの規則性を見出し、顧客の行動の予測等に活用することができる。

【0004】例えば、おむつを買った男性は、同時にビールを買う可能性が高いといった分析結果がその事例である。

【0005】このような顧客データベース中に隠された規則を抽出して意志決定支援を行う手法としてデータマイニングがあり、近年、注目を集めている。

【0006】データマイニングとは、大容量のデータに隠された因果関係やパターンを探索したりモデル化するための、高度な手法で、複雑なデータ間の関係づけを可能にするものであり、人工知能、データ工学等の分野においてここ数年注目を浴びている意思決定支援手法である。

【0007】なお、このデータマイニングは、様々な適用分野があり、例えば、過去の購買記録から考えて、顧客が次に買いそうな商品は何か、クレジットカードの不正処理がいつ起こるか、顧客がライバル会社に流れたのはなぜか、また、どうすれば取り戻せるか、或いは購買パターンによって顧客をセグメント化する最適な方法は何か、といったことについての解法の支援をしてくれる。

【0008】

【発明が解決しようとする課題】ところで、従来のデータベースは、顧客の表面的な行動をその項目として蓄積するものに過ぎず、よって、このようなデータベースにより抽出した規則により顧客のその行動の背景にある心の動きまでを追うことはできない。したがって何故その

ような行動パターンが形成されたかという理由を見つけることは一般的に困難である。

【0009】また、個人を対象としたダイレクトマーケティングをする場合でも、その人の外見的情報や行動だけで判定しており、効果的なマーケティングを行うことはできなかった。

【0010】さらに、特定の商品又はサービスに対するダイレクトマーケティングを行う場合、顧客データベース中の特定セグメント（例えば、20代女性を限定といったもの）をアプローチの対象としていたが、この方法だとセグメント中の顧客を平等に扱うため、大規模な顧客データベースからもっとも有望な顧客を順位の付けてピックアップすることは困難であった。

【0011】そして、一般的にデータマイニングは、顧客データベースの内容を分析することでデータベース中の項目間の規則を抽出するものであり、よって、顧客データベース中の項目と顧客データベース中に存在しない項目との関係を抽出することはできない。すなわち、データマイニングでは、規則を抽出する対象となる項目が明らかでない限り当該項目との間における規則を抽出することはできなく、このようなことから、既存データベースを分析しても別分野のビジネスに活用することは困難とされていた。

【0012】そこで、本発明は、上述の実情に鑑みてなされたものであって、顧客の規則抽出の精度を向上させることができる顧客情報検索方法及び顧客情報検索装置、データ作成方法、並びにデータベースの提供を目的としている。

【0013】

【課題を解決するための手段】本発明に係る顧客情報検索方法は、上述の課題を解決するために、顧客の心理側面に関する顧客内的情報と、少なくとも顧客の基本属性を含む顧客外的情報とを記憶手段に蓄積し、記憶手段に蓄積した顧客内的情報及び顧客外的情報に基づいて、顧客の特定の商品又はサービスに対する有望度を算出する。

【0014】すなわち、この顧客情報検索方法は、顧客外的情報に加え顧客内的情報に基づいて顧客の特定の商品又はサービスに対する有望度を算出する。

【0015】この顧客情報検索方法により算出された有望度は、特定の商品又はサービスの潜在顧客であるか否かを的確に示す指標となる。

【0016】また、本発明に係る顧客情報検索装置は、上述の課題を解決するために、顧客の心理側面に関する顧客内的情報と少なくとも顧客の基本属性を含む顧客外的情報とが蓄積される記憶手段と、記憶手段に蓄積した顧客内的情報及び顧客外的情報に基づいて、顧客の特定の商品又はサービスに対する有望度を算出する有望度算出手段とを備える。

【0017】このような構成を有する顧客情報検索装置

(4) 000-172697 (P2000-172697A)

は、記憶手段に蓄積されている顧客外的情報に加え顧客内的情報に基づいて、有望度算出手段により、顧客の特定の商品又はサービスに対する有望度を算出する。

【0018】この顧客情報検索装置により算出された有望度は、特定の商品又はサービスの潜在顧客である否かを的確に示す指標となる。

【0019】また、本発明に係るデータ作成方法は、上述の課題を解決するために、顧客の心理側面に関する顧客内的情報と、少なくとも顧客の基本属性を含む顧客外的情報とを記憶手段に蓄積する。

【0020】このデータ作成方法により記憶手段に記憶された顧客内的情報及び顧客外的情報を用いて算出された特定の商品又はサービスに対する有望度は、特定の商品又はサービスの潜在顧客であるか否かを的確に示す指標となる。

【0021】また、本発明に係るデータベースは、上述の課題を解決するために、顧客に関する情報から構成される既存の顧客データベースから商品又はサービスの有望度を得るために取り出した第1の顧客情報と、商品又はサービスの有望度を得るために新たに取得した顧客に関する情報を示す第2の顧客情報とを有している。そして、データベースの有する第1の顧客情報と第2の顧客情報の間に共通して存在する規則が抽出される。

【0022】このデータベースの第1の顧客情報と第2の顧客情報の間に共通して存在する規則に、顧客に関して既に得られている情報を適用して、特定の商品又はサービスの潜在顧客であるか否かの指標を得る。

【0023】また、本発明に係る顧客情報検索方法は、上述の課題を解決するために、顧客に関する情報から構成される既存の顧客データベースから商品又はサービスの有望度を得るために取り出した第1の顧客情報と商品又はサービスの有望度を得るために新たに取得した顧客に関する情報を示す第2の顧客情報の間に共通して存在する規則を抽出する。

【0024】この顧客情報検索方法により抽出した第1の顧客情報と第2の顧客情報の間に共通して存在する規則に、顧客に関して既に得られている情報を適用して、特定の商品又はサービスの潜在顧客であるか否かの指標を得る。

【0025】

【発明の実施の形態】以下、本発明の実施の形態について図面を用いて詳しく説明する。この実施の形態は、本発明を、顧客に関して蓄積された情報に基づいて、所望の顧客情報を検索する顧客情報検索装置に適用したものである。

【0026】図1に示すように、顧客情報検索装置1は、各種データの入力部とされる入力部2と、データ処理の結果等を出力する出力部3と、データを記憶するデータ記憶部4と、顧客情報検索装置1を構成する各部を制御等するCPU5とを備えている。そして、顧客情報

検索装置1は、各部がCPU5とバス6によって接続されている。

【0027】この顧客情報検索装置1において、データ記憶部4は、顧客の心理側面に関する顧客内的情報と、少なくとも顧客の基本属性を含む顧客外的情報とが蓄積される記憶手段とされ、CPU5は、データ記憶部4に蓄積した顧客内的情報及び顧客外的情報に基づいて特定の商品又はサービスに対する有望度を算出する有望度算出機能を有している。

【0028】入力部2は、データを入力するための入力手段として構成されている。例えば、入力部2は、マウス、キーボード等の入力操作手段として構成されている。

【0029】出力部3は、データを出力するための出力手段として構成されている。例えば、出力部3は、モニタ等の表示手段、又はプリンタ等の印刷出力手段として構成されている。

【0030】データ記憶部4は、各種データが記憶される記憶手段として構成されている。入力部2から入力されたデータは、このデータ記憶部4に記憶される。データ記憶部4には、顧客内的情報及び顧客外的情報等が顧客データベース等として記憶されている。このデータ記憶部4に記憶されるデータについては、後で詳しく説明する。

【0031】CPU5は、顧客情報検索装置1の各部を制御する制御手段として構成されている。そして、CPU5は、上述したように、有望度算出機能を有している。

【0032】次に、データ記憶部4に記憶されるデータについて詳しく説明する。

【0033】データ記憶部4には、顧客内的情報及び顧客外的情報が記憶されている。顧客内的情報及び顧客外的情報により、例えば後述する顧客データベースが構築されている。また、データ記録部4には、後述する学習用データベースや仮想データベースも記憶されている。

【0034】顧客外的情報は、顧客の基本属性に関する情報及び顧客の行動履歴に関する情報から構成されている。

【0035】顧客の基本属性に関する情報は、一般的にはデモグラフィック情報と呼ばれ、顧客の基本的な属性に関する情報からなり、具体的には、顧客の性別、年齢、職業、家族構成等の情報によって構成されている。

【0036】顧客の行動履歴に関する情報とは、例えば顧客の商品やサービスの購買履歴からなる情報であって、具体的には、顧客の商品又はサービスの購入、利用といった顧客の行動面の履歴情報であり、さらに、顧客がいつ、どこで、何をしたかについての情報も含まれている。例えば、顧客Aが商品Bを日時CにショップDにおいて価格Eで購入したといったことが、時系列に格納された情報がこの情報に含まれている。

(5) 000-172697 (P2000-172697A)

【0037】デモグラフィック情報と行動履歴の情報は、顧客の表面的或いは行動に現れるものであり、顧客の外的データとなる。

【0038】顧客内的情報は、顧客の価値観を少なくとも含む心的側面に関する情報であって、顧客の内的データとされる。顧客内的情報は、具体的には、テレビや雑誌といったメディアに関する考え方、生活態度、社会に対する意識、及び自身の生き方等といった情報とされる。

【0039】データ記憶部4には、上述の各種情報が、顧客データベースの顧客プロフィールデータとして記憶されている。図2には、顧客データベースDB₁が、いわゆるスプレッドシートの形態として、各行に顧客一人分の情報が格納されている例を示している。

【0040】図2の例では、デモグラフィック情報、製品／サービス購入・利用履歴、及び人柄データを、プロフィール項目としており、顧客が顧客IDによって整理されている。この図2に示す例の顧客データベースDB₁は、1000万人分のデータをファイルしている。

【0041】本例では、デモグラフィック情報が「氏名」、「電話番号」、「住所コード」、「性別」、「年齢」、「家族コード」及び「年収コード」等とされ、一部コード化された項目から構成されている。なお、この

$$F = a \times Q1 + b \times Q2 + c \times Q3 + d \times Q4 + e \times Q5 + \dots$$

... (1)

ここで、Q1、Q2、Q3、Q4、Q5、...は顧客プロフィールデータ、a、b、c、d、e、...は特定の商品又はサービス向けに得られた定数である。

【0048】なお、有望度Fの算出式はこのような線形結合式によって表現されることに限定されるものではなく、非線形のような計算式によって表現することもできる。例えば、シグモイド関数を用いたニューラルネットワークモデルがその一つである。

【0049】また、下記に示すような条件式（IF文）によるルール形式によっても有望度を導き出すことができる。

【0050】IF (Q1>a AND Q2>b AND Q3>c AND Q4>d AND Q5>e) THEN F = X

なお、線形モデルによる統計手法は、判別分析、ロジスティック回帰、クラスター分析等が該当し、「なぜ」「どのように」といった疑問の返答に適している。

【0051】また、非線形モデルの一つであるツリーモデル（帰納方式とも呼ばれる）は、データからディビジョンツリー（決定木）を構成している。このモデルは、重要な変数を選び、不要な予測要素を削除する場合に適している。

【0052】そして、非線形モデルであるニューラルネットワークは、履歴データをもとに未来の結果を予測することができ、「何が」という疑問の返答に適している。

デモグラフィック情報は、従来のデータベースでも採用されている基本属性であり、いわゆる名寄せ情報にも使われている情報である。

【0042】また、製品／サービス購入・利用履歴は、行動履歴に対応するもので、「製品コード」、「購入日時」及び「製品コード」等とされ、一部コード化された項目から構成されている。

【0043】そして、人柄データは、顧客内的情報に対応されるものであり、「質問1」、「質問2」、「質問3」、...というように各種質問の回答情報により構成することが作成方法の一例である。

【0044】顧客情報検索装置1は、このように構成された顧客データベースDB₁から特定の商品又はサービスに対する有望顧客のピックアップを行うことができる。特定の商品又はサービスに対する顧客の有望度について説明する。

【0045】有望度は、特定の商品又はサービスに対する顧客の関心や需要等を示す指標となるもので、例えば計算式などによる規則抽出により求めることができる。

【0046】有望度Fの算出式には、簡易なものとして、数値化されたデータ項目に係数を掛け合わせた(1)式のように示される線形結合式がある。

【0047】

【0053】顧客情報検索装置1は、このような計算式等により、データマイニング手法を採用して、特定の商品又はサービスに対する各顧客の有望度を算出している。

【0054】顧客情報検索装置1は、特定の商品又はサービスについて顧客の各々に対し求めた有望度に基づいて、その測定の商品又はサービスの潜在顧客をピックアップしている。例えば、これにより、効率良くマーケティングを行うことができる。すなわち、外的データに加えて心理的要因をあらわす内的データを組み合わせること、つまり、顧客の表面的あるいは行動にあらわれるデータと内面的な人柄のデータを組み合わせることで個人を完全に把握することができ、効率のいいマーケティングを行うことができる。例えば、特定の商品又はサービスの購入を望む顧客に対してダイレクトメール等を送付するなどといったことが可能になる。

【0055】また、有望度の算出式は、任意の方法で作成することもできる。例えば、顧客から得た各種情報に基づいて、その情報間に存在する関係から有望度の算出式を作成することもできる。一般には顧客のプロファイルデータと予測したい顧客の嗜好等の関係に対応する変数（ターゲット変数）が既知であるデータセット（学習用データベース）に対して、上述のモデルを適用して各種のパラメータを獲得することができる。

【0056】学習データベースは、例えば図3中(A)

(6) 000-172697 (P2000-172697A)

及び(B)に示すように、顧客データベース(図3中(A))から任意にサンプリングして作成したデータベース(図3中(B))とされていた。従来は、このように顧客データベースの一部項目のみからなる学習用データベースに基づいて、その顧客に対するダイレクトメール等のレスポンスデータを使って予測していた。

【0057】しかし、本発明を適用した学習用データベースは、顧客データベースの情報から作成するなどせずに全く別の手段により収集したデータにより構成している。すなわち、例えば、学習用データベースは、ランダム、層別等で抽出した適当な人数、例えば1000~10000人程度にアンケート調査をすること等により顧客データベースとは全く関係のない項目を有するデータベースとして作成される。図4には、顧客情報検索装置1が使用する学習用データベースDB₂の構成例を示している。

【0058】図4に示すように、学習用データベースDB₂は、アンケート項目が大きく二つの部分からなり、図2に示した顧客データベースDB₁と同一項目、例えばデモグラフィック情報、製品/サービス購入・利用履歴や人柄データ等の項目からなるデータベースDB₂₁の部分と、顧客データベースに格納されていない項目であって、潜在顧客の判定等に用いるターゲット変数の項目からなるデータベースDB₂₂の部分とからなる。例えば、データベースDB₂₂の項目は潜在顧客判定のためにした質問の項目からなり、具体的には、趣味や嗜好等の様々な商品又はサービスに対する購入希望、或いは心理的データといったダイレクトマーケティング項目からなる。

【0059】このように通常業務の顧客データベースから得ることができない情報を適当な抽出によりアンケートして得た情報等により学習用データベースを作成することで、顧客データベースに欠けている様々な項目を推定することが可能になる。

【0060】図5には、上述したような学習用データベースを含んで構成されるデータベースを示している。図5に示すように、データベースは、顧客データベースDB₁、学習用データベースDB₂及び仮想データベースDB₁₂より構成されている。

【0061】ここで、顧客データベースDB₁は、例えば図2に示したデータベースの項目からなり、顧客データベースと学習データベースの共通項目とからなるデータ項目Q₁、・・・からなる。また、学習用データベースDB₂は、例えば図4に示した学習用データベースDB₂である。

【0062】仮想のデータベースDB₁₂は、学習用データベースDB₂のみが有している学習データベース単独項目からデータ項目Q_x、・・・からなる。この仮想データベースは、学習モデルを適用して、データベースDB₂₁の項目とデータベースDB₂₂の項目との規則抽出に

よる定式化により推定された情報からなり、すなわち、顧客から実際に得ていないデータ項目Q_x、・・・の情報が推定されたデータベースを構築している。

【0063】ここで、データベースDB₂₁の項目とデータベースDB₂₂の項目の間の定式化を可能にする学習用データベースに適用される学習モデルとして、線形判別式の場合には判別分析等を用い、ニューラルネットワークの場合にはバックプロパゲーションアルゴリズムを用い、ルール形式の場合には判別木による帰納的解法を用いる。

【0064】顧客の持つ様々な趣味、嗜好、或いは価値観といったものを顧客に直接たずねることは、多くの場合、コスト、プライバシーの問題等で困難である。しかし、このように、学習用データベースを用いて顧客データベースに欠けている様々な項目をについて情報を推定し、仮想データベースを構築することで、実質的に低コストで巨大なデータベースを持つことと同じ効果を得ることができる。

【0065】また、学習用データベースは、新商品や新サービスのダイレクトマーケティング等の様々な目的に応じて随時、アンケート調査等で拡張することが可能である。

【0066】次に、顧客データベース中から特定の商品又はサービスの潜在顧客を抽出する場合について図6を用いて具体的に説明する。この例では、音楽専門チャンネルの潜在顧客を抽出する場合を示しており、顧客データベースDB₁から2段階に分けて音楽専門チャンネルの潜在顧客DB₄をピックアップしている。

【0067】最初の段階では、購買履歴や性別、年齢といった顧客データベースDB₁から得られる情報により、顧客についての最初の絞り込み、すなわち、プリサンプリングを行う。これにより、分析顧客母集団からなるデータベースDB₃が作成される。例えば、顧客データベースDB₁は1000万人規模からなるデータベースであり、この絞り込みにより、数百万規模の分析顧客集団からなるデータベースDB₃が作成される。

【0068】例えば、分析顧客母集団のデータベースDB₃は、「30代既婚」、「大型テレビ購入」、「ステレオコンボ購入」、及び「新製品早期購入傾向」といった項目に該当する顧客に絞り込まれる。例えば、従来のダイレクトマーケティングでは、この段階までの顧客の抽出が可能とされていた。

【0069】次に、学習用のデータベースDB₂を用いて顧客の様々な特性を予測することにより、分析顧客母集団のデータベースDB₃から絞り込まれた音楽専門チャンネルの潜在顧客のデータベースDB₄が作成される。ここで、学習用データベースDB₂は例えば1000人規模のデータベースからなる。

【0070】上述したような学習用データベースDB₂を用いた絞り込みにより、例えば、音楽専門チャンネル

(7) 000-172697 (P2000-172697A)

の潜在顧客のデータベースDB₄は、「テレビは娯楽として手軽」、「音楽のない生活は考えられない」、「ワクワクするような刺激を求める」、「音楽専門チャンネル利用意向」、及び「ポップスが好き」等といった項目に該当する顧客によって構成されるようになる。

【0071】以上のように、顧客情報検索装置1は、顧客データベースからなる分析顧客母集団のデータベースに、学習用データベースの情報を適用させて、特定の商品又はサービスの潜在顧客の抽出を行っている。

【0072】なお、この例では絞り込みを2段階に分けておりその段階を踏む順序については、上述した例と逆の順番にしたり、或いは一段階目を省略することもできる。

【0073】次に、学習用データベースにおける処理と顧客データベースにおける処理の手順について図7を用いて具体的に説明する。ここにおける処理は、例えばCPU5により実行される処理である。

【0074】この図7に示すように、ステップS1で処理が開始された学習用データベースでは、ステップS2に示すように、CPU5により、予測したい顧客の特性を学習用データベースの項目から選択する。この項目は新商品及びサービスの場合、通常、顧客データベースには含まれないものである。この選択した項目が学習のターゲット変数になる。

【0075】次に、ステップS2において、CPU5は、顧客の特性を予測する変数となる独立変数を作成する。例えば、そのために学習用データベースと顧客データベースとの共通項目からプロフィール項目を抽出する。

【0076】そして、CPU5は、ステップS4において、この選ばれたプロフィール項目に対して、データの抜け、異常値等に対する処理等を行う欠損値処理をして、続くステップS5において、プロフィール項目について編集、合成等の処理等を行う変数加工をして、最終的な独立変数を作成する。

【0077】続いて、CPU5は、ステップS6及びステップS14において、学習用データベースを規則発見用のデータ（学習用データ）及び抽出規則の評価用データに分割する。

【0078】規則発見用のデータにより、CPU5は、ステップS7において、学習モデルを選択して、判別分析（ステップS8）、デシジョンツリー（ステップS9）、ニューラルネットワーク（ステップS10）等の様々な学習モデルを適用させる。

【0079】このように各学習モデルを適用した後、CPU5は、ステップS11において、評価用データとその学習モデルの適用結果に基づいて、もっとも有効な学習モデルを選択し、ステップS12において、選択した学習モデルにより規則のパラメータを決定し、ステップS13において、その規則のパラメータを保存する。

【0080】一方、顧客データベース側でも学習用データベースで作成した独立変数と同一のものを同一の手順により作成する。これにより学習用データベースで作成された規則は顧客データベースでもまったく適用可能になる。

【0081】具体的には、顧客データベースに関しては、ステップS21において処理が開始され、ステップS22において、名寄せ処理が行われる。続いて、CPU5は、学習データベースにおける処理と同様に、ステップS23～ステップS25において、独立変数の作成、欠損値処理、及び変数加工といった処理を行う。

【0082】この各種処理を終えた後、CPU5は、ステップS26において、顧客データベース中からデモグラフィック情報、購買履歴情報等に基づいて顧客の絞り込み、すなわち、顧客のプリサンプリングを行う。

【0083】CPU5は、ステップS26により顧客の絞り込みを行った後、ステップS27において、学習用データベース処理のステップS13において得た学習モデルを適用する。

【0084】この学習モデルの適用により、CPU5は、ステップS28及びステップS29において、ターゲットの顧客、すなわち、潜在顧客を抽出して、抽出した顧客（抽出顧客）に抽出済のフラグ付けを行う。そして、CPU5は、ステップS30において、名簿等として抽出した顧客の出力を行う。名簿等としての出力は、出力部3を構成するモニタ又はプリンタ等によって行われる。

【0085】以上が顧客情報検索装置1の構成と、その処理の手順の説明である。顧客情報検索装置1は、購買履歴等の行動履歴等からなる顧客外的情報に加え、顧客の価値観等の心理側面に関する情報からなる顧客内的情報を使用し、さらには顧客から直接得ることができない情報を推定することにより、顧客の規則抽出の精度を向上させている。

【0086】すなわち、例えば、顧客情報検索装置1は、図8に示すように、性別、年齢及び家族等のデモグラフィック情報、並びにAV機器購入履歴及びサービス購入履歴等の行動履歴などからなる外的データ（図8中（A））に加え、メディアに関する意識、及びライフスタイル特性等の人柄に関する情報等からなる内的データ（図8中（B））をも参照することにより、規則抽出の精度を向上させ、所望の顧客情報を抽出している。例えば、これにより、個々の顧客に最適なコンテンツ、サービス及びプロダクツの提供（図8中（C））が可能になる。

【0087】また、顧客情報検索装置1は、データベース中の顧客を特定の商品又はサービスに対する有望度の大きさによって順位づけすることもでき、これにより、有望度の上位の順からマーケティングに必要な人数だけ、或いは有望度が所定の閾値以上になっている顧客だ

(8) 000-172697 (P2000-172697A)

けをピックアップしてダイレクトアプローチすることが可能になり、効率的なマーケティングが実現できるようになる。

【0088】例えば、ある商品又はサービスの購入希望者をピックアップすることを想定する。有望度算出式を購入希望者と購入非希望者の線形判別関数として求めた場合の有望度の分布が図9に示すようになったとする。例えば、商品又はサービスの購入希望者による関数 f_1 は、音楽専門チャンネルをぜひ利用したい顧客の集合から求めたものであり、購入非希望者による関数 f_2 は、音楽専門チャンネルを利用したくない顧客の集合から求めたものである。

【0089】この場合、閾値1近辺の有望度をもつ顧客を選んでも購入希望に対する的中率はほぼ0%となる。しかし閾値2近辺の有望度を持つ顧客を選んだ場合は的中率はほぼ50%に、そして閾値3以上の顧客を選んだ場合は的中率がほぼ100%となる。このような中率の分布を想定することで、大規模な顧客データベース中から必要人数だけの潜在顧客をピックアップすることが可能になる。例えば、100万人の顧客データベースから有望度の高い上位1万人を抽出した際の的中率が90%であるとき、1万人にダイレクトメールを出す場合には9千人が的中した潜在顧客になる。言い換えると9千人の顧客が欲しい場合には1万人にダイレクトメールを出せば良いことになる。

【0090】さらに、学習データベースを用いて顧客の情報を推論することは、潜在顧客を適切に抽出することを可能にし、例えば図10に示すような手順により、マーケティングを行うことを可能にする。

【0091】顧客プロフィール情報を選択し(ステップS31)、その顧客プロフィール情報に基づいて推論を行う(ステップS32)。ここで、推論は、例えばデータマイニング手法を採用している。

【0092】そして、この人にはこのような方法で売り込めば成功する(ステップS33)といった流れを得る。

【0093】このように、学習用データベースにより得た規則を適用し、顧客の情報を推論することにより、潜在顧客を適切に抽出することが可能になり、マーケティング戦略をより効果のあるものとすることが可能になる。

【0094】すなわち、顧客データベース中のデータ項目と顧客データベースには存在しないデータ項目の規則を学習用のデータベースを用いて抽出することで、既存のデータベースを新規ビジネスに活用可能にすることが可能になる。さらに、顧客に直接質問することなく、必要に応じて学習用データベースを用いて顧客データベースには存在しないデータ項目を推定することで、プライバシーを侵害することなく、低コストでまた記憶装置の容量を拡大することなく実質上のデータベースを拡大す

ることができる。

【0095】次に、コンピュータシミュレーションにより、音楽専門チャンネルの潜在顧客発見を題材にして、潜在顧客の抽出的中率を測定した実施例について説明する。コンピュータシミュレーションは、線形モデルの判別分析を採用して行っている。なお、線形モデルによりコンピュータシミュレーションを行うことに限定されず、ディビジョンツリー等の他の学習モデルを採用してコンピュータシミュレーションすることとしても良い。

【0096】図11に示すように、顧客データベースD_{B1}からプロファイリングデータを用いて特定の商品又はサービスの潜在顧客を検索した場合に、抽出された潜在顧客からなるデータベースD_{B4}において確かな潜在顧客とされる的中率が問題になる。実施例では、この的中率を中心に説明する。

【0097】実施例において、パラメータを次のように設定している。

【0098】プロフィールデータとして、例えば表1に示すように、AV機器購入履歴、デモグラフィック、人柄データ及びレンタルビデオ・CD購入頻度の各情報を用いている。

【0099】

【表1】

プロフィール情報
AV機器購入履歴 デモグラフィック 人柄データ レンタルビデオ・CD購入頻度

【0100】AV機器購入履歴については、例えば表2に示すように、横長ワイドテレビ、ハイビジョンテレビ、・・・等のようなパラメータを使用している。

【0101】

【表2】

1. 横長ワイドテレビ
2. ハイビジョンテレビ
3. ビデオプロジェクター
4. 小型液晶カラーテレビ
5. カラーテレビ14型以下
6. カラーテレビ15～21型
7. カラーテレビ22～29型
8. カラーテレビ30～34型
9. カラーテレビ35型以上
10. 衛星放送用アンテナ

(9) 000-172697 (P2000-172697A)

【0102】人柄データについては、例えば表3に示すように、情報通信に関する考え方1として、自分の目では見えないものを見たい、労働時間が短縮され自分の時間や生活のゆとりがもてる等のようなパラメータを使用

している。

【0103】

【表3】

情報通信に関する考え方1

- ・自分の目では見えないものを見たい
- ・労働時間が短縮され自分の時間や生活のゆとりが持てる

情報通信に関する考え方2

- ・お金がかかっても自分の関心事は詳しく知りたい
- ・自分なりの情報の集め方をもっているほうだ
- ・社会の情報化に取り残されたくないと思う
- ・仲間内の口コミ情報を重視するほうだ

【0104】図12には、上述したようなパラメータを用いたコンピュータシミュレーションにより得た結果を示して。図12のグラフの横軸は、潜在顧客のデータベースから抽出した割合であって、可能性の高い人から順に示されている。すなわち、0%に近いほど、音楽専門チャンネルの潜在顧客であろう顧客になる。グラフの縦軸は、的中率を示している。

【0105】この図12に示す結果において、確からしい10%の人を選んだ場合(値 X_1)には、90%の確立で潜在顧客を抽出することができる。そして、確からしい5%の人を選んだ場合(値 X_2)には、ほぼ100パーセントの確立で潜在顧客を抽出することができる。

【0106】なお、無作為に人を選んだ場合(値 X_3 (100%))には、25%の確立で潜在顧客を抽出することができる。

【0107】図13は、比較例を示している。図12に示した本発明を適用した場合の的中率の変化 f_{11} に対比して、人柄データのみから得られる的中率の変化 f_{12} 及びCD・ビデオレンタル回数及びデモグラフィックから得られる的中率の変化 f_{13} の結果を示している。なお、図13に示すグラフの横軸は、抽出した人数であり、少ない数になるほど音楽専門チャンネルの潜在顧客であろう顧客になる。

【0108】この図13に示すように、CD・ビデオレンタル回数が多い顧客、すなわちCD・ビデオレンタルのヘビーユーザを抽出した場合(f_{13} において抽出人を1人目に近づけた場合)であっても、的中率は低いままであり、本発明を適用した学習モデルにより得た的中率との間にかなりの差が生じていることがわかる。

【0109】なお、このように正確に潜在顧客を抽出した要因として、例えば、エンタテインメントを主眼とした自由度の高いプロフィールデータを採用していること、大まかに傾向をみるクラスタアプローチではなく、個人を分析するパーソナルアプローチを採用しているこ

と等が挙げられる。

【0110】以上がコンピュータシミュレーションによる実施例の結果でも、本発明を適用することにより、的中率を上げることができ、ほぼ100%の確率で潜在顧客を抽出することが可能である結果を得ることができた。顧客情報検索装置1は、このように、特定の商品又はサービスの潜在顧客を的確に抽出することを可能にする。

【0111】

【発明の効果】本発明に係る顧客情報検索方法は、顧客の心理側面に関する顧客内的情報と、少なくとも顧客の基本属性を含む顧客外的情報とを記憶手段に蓄積し、記憶手段に蓄積した顧客内的情報及び顧客外的情報に基づいて、顧客の特定の商品又はサービスに対する有望度を算出することにより、潜在顧客の規則抽出の精度を向上させることができる。

【0112】よって、顧客情報検索方法により、特定の商品又はサービスの潜在顧客を的確に抽出することが可能になる。

【0113】また、本発明に係る顧客情報検索装置は、顧客の心理側面に関する顧客内的情報と少なくとも顧客の基本属性を含む顧客外的情報とが蓄積される記憶手段と、記憶手段に蓄積した顧客内的情報及び顧客外的情報に基づいて、顧客の特定の商品又はサービスに対する有望度を算出する有望度算出手段とを備えることにより、記憶手段に蓄積されている顧客内的情報及び顧客外的情報に基づいて、有望度算出手段により、顧客の特定の商品又はサービスに対する有望度を算出することができる。

【0114】これにより、顧客情報検索装置は、潜在顧客の規則抽出の精度を向上させることができ、特定の商品又はサービスの潜在顧客を的確に抽出することを可能にする。

【0115】また、本発明に係るデータ作成方法は、顧

(10) 00-172697 (P2000-172697A)

客の心理側面に関する顧客内的情報と、少なくとも顧客の基本属性を含む顧客外的情報とを記憶手段に蓄積することにより、記憶手段に記憶された顧客内的情報及び顧客外的情報を用いた顧客の規則抽出を可能にする。そして、この顧客内的情報及び顧客外的情報順に基づいて行う潜在顧客の規則抽出の精度は高いものとなる。

【0116】よって、このデータ作成方法により、特定の商品又はサービスの潜在顧客を的確に抽出することが可能になる。

【0117】また、本発明に係るデータベースは、顧客に関する情報から構成される既存の顧客データベースから商品又はサービスの有望度を得るために取り出した第1の顧客情報と、商品又はサービスの有望度を得るために新たに取得した顧客に関する情報を示す第2の顧客情報とを有することにより、第1の顧客情報と第2の顧客情報の間に共通して存在する規則が抽出することができ

る。

【0118】この抽出した規則に、顧客に関して既に得られている情報を適用することにより、例えば、特定の商品又はサービスの潜在顧客を的確に抽出することが可能になる。

【0119】また、本発明に係る顧客情報検索方法は、顧客に関する情報から構成される既存の顧客データベースから商品又はサービスの有望度を得るために取り出した第1の顧客情報と商品又はサービスの有望度を得るために新たに取得した顧客に関する情報を示す第2の顧客情報の間に共通して存在する規則を抽出することにより、この抽出した規則を、顧客に関して既に得られている情報を適用することによってできる。これにより、例えば、特定の商品又はサービスの潜在顧客を的確に抽出することが可能になる。

【図面の簡単な説明】

【図1】本発明の実施の形態である顧客情報検索装置の構成を示すブロック図である。

【図2】本発明を適用して構成した顧客データベースの構成を示す図である。

【図3】顧客データベースから学習用データベースを作成する場合の説明に用いた図である。

【図4】本発明を適用して構成した学習用データベースの構成を示す図である。

【図5】顧客データベースと学習用データベースとからなるデータベースの構成を示す図である。

【図6】顧客データベースから得た分析顧客母集団に学習用データベースを適用して潜在顧客を抽出する場合の手順を示す図である。

【図7】学習用データベースに基づいて行う処理の手順と、顧客データベースに基づいて行う処理の手順とを示す図である。

【図8】内的データと外的データとから特定の商品又はサービスの潜在顧客を抽出する場合をモデル化して示す図である。

【図9】顧客の有望度と潜在顧客の的中率との関係を示す特性図である。

【図10】顧客情報検索装置により可能になるマーケティングを説明するために用いた図である。

【図11】顧客データベースから潜在顧客からなるデータベースを作成した場合の的中率を説明するために用いた図である。

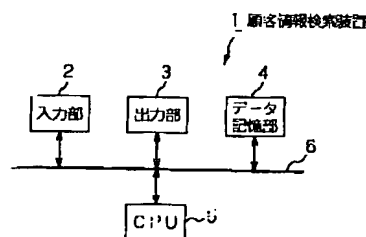
【図12】抽出した潜在顧客と、確かな潜在顧客である的中率との関係を示すものであって、本発明を適用した実施例の結果を示す特性図である。

【図13】抽出した潜在顧客と、確かな潜在顧客である的中率との関係を示すものであって、本発明を適用した実施例の結果と従来の結果の比較例を示す特性図である。

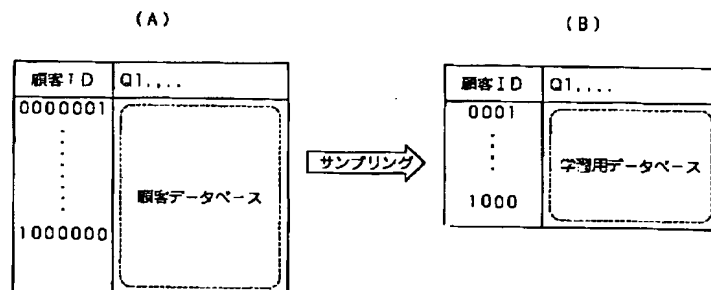
【符号の説明】

1 顧客情報検索装置、4 データ記憶部、5 CPU

【図1】



【図3】



(表 1) 00-172697 (P2000-172697A)

【図 2】

プロフィール 項目	デモグラフィック情報							製品/サービス 購入・利用履歴				人柄データ							
	氏名	電話番号	住所 コード	性別	年齢	家族 コード	年収 コード	製品 コード	購入 日時	製品 コード	...	質問 1	質問 2	質問 3	...	質問 17	質問 18
顧客ID																			
100.000.001	安藤	2544	27	1	35	2	3	3	1999.4.01	7	...	5	3	4	...	2	1
100.000.002	田中	2878	32	1	24	4	4	7	1999.4.03	6	...	4	2	1	...	3	4
100.000.003	山下	1112	14	2	47	1	6	7	1999.4.15	4	...	5	1	1	...	5	4
100.000.004	鈴木	7847	41	1	18	3	0	2	1999.5.17	4	...	2	1	3	...	1	5
100.000.005	斎藤	6245	22	2	33	3	2	1	1999.5.08	8	...	5	1	5	...	3	3
...
...
109.999.998	山田	1383	27	2	42	4	3	7	2001.7.17	5	...	1	2	1	...	3	2
109.999.999	清水	2998	13	1	51	2	8	7	2001.8.27	4	...	2	4	5	...	2	4
110.000.000	川崎	6555	12	1	21	1	2	6	2001.9.12	2	...	5	5	4	...	4	5

DB1 顧客データベース

【図 4】

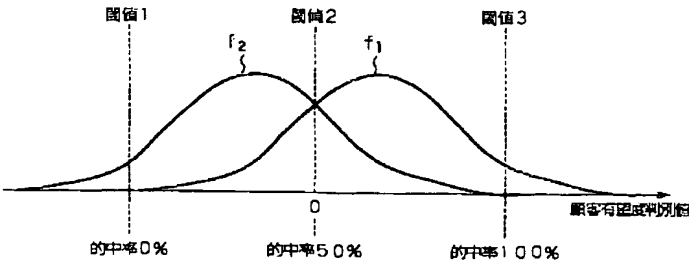
プロフィール 項目	デモグラフィック 情報							製品/サービス 購入・利用履歴				人柄データ								趣味嗜好データ		
	住所 コード	性別	年齢	家族 コード	年収 コード	製品 コード	購入 日時	製品 コード	...	質問 1	質問 2	質問 3	...	質問 17	質問 18	...	質問 1	質問 2	質問 3
顧客ID																						
001	27	1	35	2	3	3	1999.4.01	7	...	5	3	4	...	2	1	...	5	3	4
002	32	1	24	4	4	7	1999.4.03	6	...	4	2	1	...	3	4	...	4	2	1
003	14	2	47	1	6	7	1999.4.15	4	...	5	1	1	...	5	4	...	5	1	1
004	41	1	18	3	0	2	1999.5.17	4	...	2	1	3	...	1	5	...	2	1	3
005	22	2	33	3	2	1	1999.5.08	8	...	5	1	5	...	3	3	...	5	1	5
...
...
...
998	27	2	42	4	3	7	2001.7.17	5	...	1	2	1	...	3	2	...	1	2	1
999	13	1	51	2	8	7	2001.8.27	4	...	2	4	5	...	2	4	...	2	4	5
1000	12	1	21	1	2	6	2001.9.12	2	...	5	5	4	...	4	5	...	5	5	4

DB21

DB22

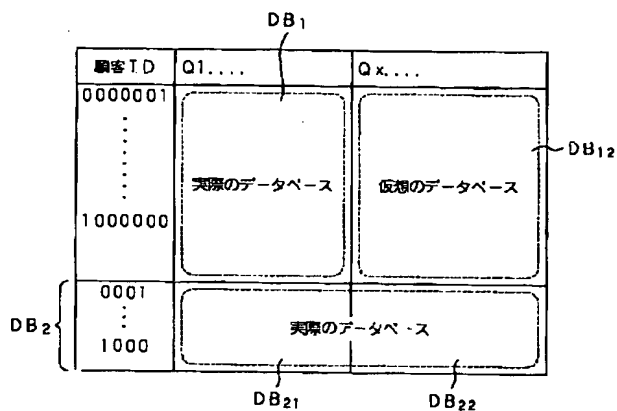
DB2 学習用データベース

【図 9】

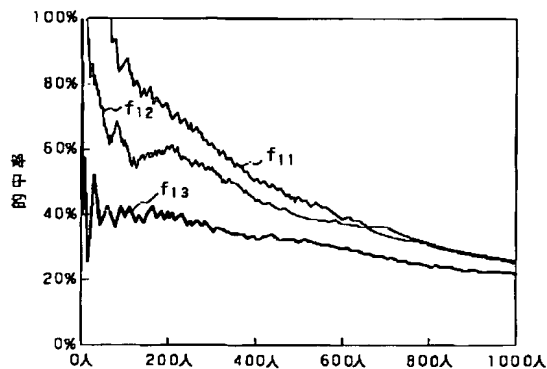


(12) 100-172697 (P2000-172697A)

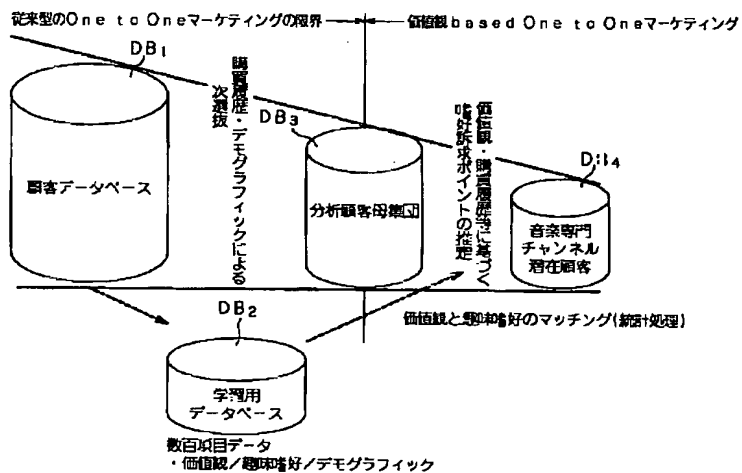
【図5】



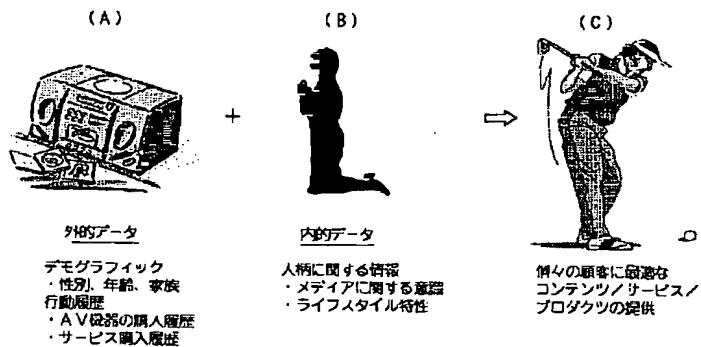
【図13】



【図6】

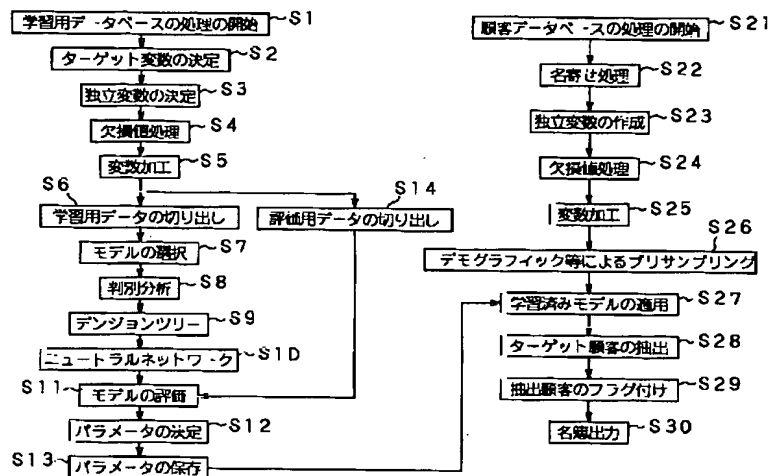


【図8】

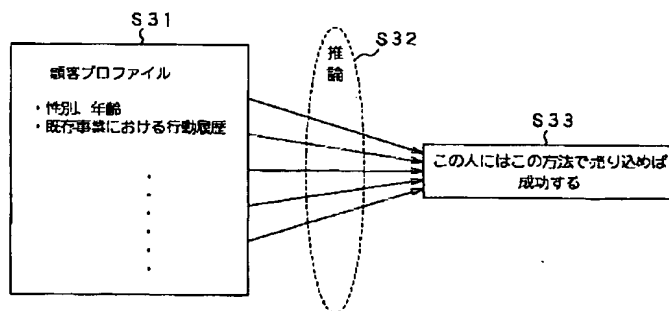


(13) 00-172697 (P2000-172697A)

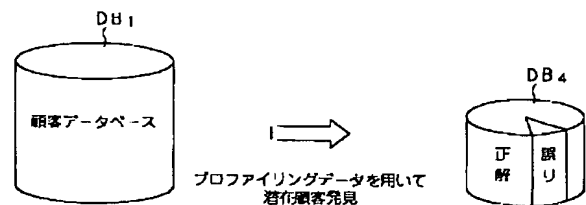
【図7】



【図10】



【図11】



(図4) 00-172697 (P2000-172697A)

【図12】

